

SYNTHETIC GENES

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to provisional application Ser. No. 61/475,921 filed Apr. 15, 2011, the disclosure of which is incorporated herein by reference in its entirety.

BACKGROUND OF THE INVENTION

[0002] To achieve desired expression levels of heterologous proteins in transgenic plants it has been found beneficial to alter the native, sometimes referred to as wild-type or original, DNA coding sequence in various ways, e.g. so that codon usage more closely matches the codon usage of the host plant species, and/or so the G+C content of the coding sequence more closely matches the G+C level typically found in coding sequences of the host plant species, and/or so that certain sequences that destabilize mRNA are removed. Expression in plants of *Bacillus thuringiensis* (B.t.) crystal protein insect toxins, for example, has been improved using one or more of these approaches. See, for example, U.S. Pat. Nos. 5,380,301, 5,625,136, 6,218,188, 6,340,593, 6,673,990, 7,741,118. Codon degeneracy allows one to make synthetic DNA sequences that encode a protein of interest using codons that differ from those used in the original DNA coding sequence.

[0003] In regard to removing sequences that may destabilize mRNA, U.S. Pat. No. 7,741,118 discloses a list of 16 polyadenylation signal sequences (column 15, Table II) and calls for reducing the number of such sequences in synthetic coding sequences that are intended for expression in plants. The polyadenylation signal sequences listed in U.S. Pat. No. 7,741,118, Table II are listed below in Table 1:

TABLE 1

| Polyadenylation signal sequences listed in U.S. Pat. No. 7,741,118, Table 11. | |
|---|--------|
| 1 | AATAAA |
| 2 | AATAAT |
| 3 | AACCAA |
| 4 | ATATAA |
| 5 | AATCAA |
| 6 | ATACTA |
| 7 | ATAAAA |
| 8 | ATGAAA |
| 9 | AAGCAT |
| 10 | ATTAAT |
| 11 | ATACAT |
| 12 | AAAATA |
| 13 | ATTAAA |
| 14 | AATTAA |

TABLE 1 -continued

| Polyadenylation signal sequences listed in U.S. Pat. No. 7,741,118, Table 11. | |
|---|--------|
| 15 | AATACA |
| 16 | CATAAA |

[0004] U.S. Pat. No. 7,741,118 also calls for preferably removing the sequence ATTTA (known as the Shaw-Kamen sequence), because it has been identified as potentially destabilizing mRNA.

[0005] Contrary to the teaching of U.S. Pat. No. 7,741,118, we have found that reduction in the number of the polyadenylation signal sequences identified in Table 1 above is neither necessary nor sufficient to enable enhanced expression of synthetic genes in plants.

SUMMARY OF THE INVENTION

[0006] Table 2 below identifies 20 potential polyadenylation signal sequences that occur frequently in maize genes.

TABLE 2

| Potential polyadenylation signal sequences found in maize genes | |
|---|--------|
| 1 | ATATAT |
| 2 | TTGTTT |
| 3 | TTTTGT |
| 4 | TGTTTT |
| 5 | TATATA |
| 6 | TATTTT |
| 7 | TTTTTT |
| 8 | ATTTTT |
| 9 | TTATTT |
| 10 | TTTATT |
| 11 | TAATAA |
| 12 | ATTTAT |
| 13 | TATATT |
| 14 | TTTTAT |
| 15 | ATATTT |
| 16 | TATTAT |
| 17 | TGTTTG |
| 18 | TTATAT |
| 19 | TGTAAT |
| 20 | AAATAA |